

# Design and Analysis of Ranavirus Studies: Surveillance and Assessing Risk

Matthew J. Gray, Jesse L. Brunner, Julia E. Earl, and Ellen Ariel

## 1 Introduction

There is mounting evidence that ranaviruses can impact populations of ectothermic vertebrate species, and may contribute to species declines (Teacher et al. 2010; Earl and Gray 2014; Price et al. 2014). Studies can be designed to determine the distribution and prevalence of ranavirus, the risk of introducing the pathogen into an uninfected area, and its possible effects on populations. Properly designed studies rely on a combination of field data, laboratory experiments, and quantitative analyses, which typically require teams of experts with adequate resources. The financial cost to assess the risk of ranaviruses can be substantial. For example, Project RANA (Risk assessment of new and emerging systemic iridoviral diseases for European fish and aquatic ecosystems) cost approximately 1.4 M € (in 2012; Evira 2013). Similarly, the Maryland Department of Natural Resources (MDNR) is currently performing a surveillance study across seven US states for \$178,000 (in 2014 USD; Smith et al. 2014). According to the Global Ranavirus Consortium (GRC) website, average cost of genomic DNA (gDNA) extraction and quantitative PCR to test for ranavirus is about \$25 USD (in 2014)

---

M.J. Gray (✉)

Center for Wildlife Health, Department of Forestry, Wildlife and Fisheries, University of Tennessee, 274 Ellington Plant Sciences Building, Knoxville, TN 37996-4563, USA  
e-mail: [mgray11@utk.edu](mailto:mgray11@utk.edu)

J.L. Brunner

School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA

J.E. Earl

National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN 37996-3410, USA

E. Ariel

College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD 4811, Australia

per sample (<http://www.ranavirus.org/>). Considering that adequate sample sizes to detect ranavirus and obtain precise estimates of infection prevalence can be high ( $n > 60$ ), laboratory expenses associated with ranavirus testing are substantial. Costs for mobilizing field crews over large geographic regions are also considerable. For example, over 95% of the MDNR's budget for the above study was dedicated to personnel, field supplies, and travel. Thus, organizations that are interested in assessing the risk of ranaviruses in wild and captive populations should be prepared to invest adequate resources. Significant planning also is essential to ensure that sufficient sample sizes are collected, contamination of samples is minimized (Miller et al. 2015), and that the information collected leads to intended measurable deliverables. For organizations that have limited knowledge about ranaviruses in their region, it may take several years to document the distribution of ranaviruses, identify infection hotspots, and implement disease intervention strategies that thwart the introduction of ranavirus or reduce its prevalence. This chapter provides the basics for designing studies to assess the risk of ranavirus. In addition, we encourage organizations to collaborate with experts that have been studying ranaviruses. The GRC can provide information on ranavirus experts in your region.

## 2 Ranavirus Surveillance

The emergence of infectious diseases has mobilized universities and organizations to determine the risk of pathogens in wild populations. To quantify risk, a fundamental understanding of the host–pathogen system at molecular (Jancovich et al. 2015a) and organismal levels (Brunner et al. 2015) is essential. The assessment of risk often starts with determining whether a pathogen is emerging, which means the pathogen is increasing in geographic distribution, prevalence in a population, or host range (Wobeser 2006). In this chapter, we refer to an outbreak as an increase in ranavirus occurrence beyond background levels, which are often unknown. Because estimates of infection prevalence and incidence are used to make decisions about risk, pathogen surveillance programs are commonly employed. If designed properly, surveillance programs can be effective at detecting pathogens, obtaining precise estimates of prevalence and incidence, and providing the necessary data to determine if a pathogen is a threat to a population or species.

### 2.1 *Interpreting Infection Data*

Increasingly, ranavirus infections are detected using PCR-based methods, but other methods are also important for directly detecting the virus (i.e., isolation in cell culture, electron microscopy, antigen capture ELISA) or evidence of infection

(e.g., histology, serologic methods; Miller et al. 2015). Estimates of viral load via quantitative PCR or cell culture-based methods (i.e., plaque assays or 50 % tissue culture infective dose [TCID<sub>50</sub>]), along with other diagnostic tools (e.g., histology), can provide information on the intensity or severity of infection and disease (Miller et al. 2015). It is important to understand what each of these methods detects (e.g., PCR detects the presence of ranavirus DNA while isolation of a virus with cell culture demonstrates the presence of infectious virions, Miller et al. 2015), as well as to recognize the limitations of each. The ability to detect an infection generally increases with time since pathogen exposure, severity of the infection, and the sensitivity of the method (Miller et al. 2015). If the assay's sensitivity and specificity are known, these values should be used to adjust estimates of prevalence and incidence; if they are not known, one should interpret infection data conservatively.

The most common variables measured during surveillance studies are infection prevalence and incidence. Infection prevalence is the number of individuals infected divided by the sample size, and it estimates the proportion of the population that is infected at a particular time. A related variable is seroprevalence, which estimates the proportion of the population that has serologic evidence of prior exposure to the pathogen. Incidence, on the other hand, is the rate at which individuals become infected over a specified time period (Wobeser 2006). While it is often expressed as the number of new cases per unit time, it is generally more useful to present as per capita incidence (e.g., per 1,000 individuals at risk).

While prevalence is more commonly estimated than incidence during surveillance studies, it is simply a “snap shot” of the infection burden at a given time so it is difficult to interpret in the absence of biological context. An understanding of the relative susceptibility of species to ranavirus can help interpret prevalence data. For example, if experimental exposures show that a species dies rapidly following ranavirus exposure, then high prevalence would be most consistent with sampling during the peak of an epidemic. Several studies have reported species-level susceptibility under controlled conditions (e.g., Hoverman et al. 2011; Brenes et al. 2014b; Brunner et al. 2015). Biological context also can be gleaned from the density of the population and the timing of the survey relative to the phenology of the organism. For example, observing low prevalence and a dense population of amphibian larvae early in spring would be more consistent with the virus recently being introduced rather than an outbreak already occurring.

Several surveillance studies suggest background prevalence levels for ranavirus in amphibian and chelonian populations is <5 % (e.g., Todd-Thompson 2010; Hoverman et al. 2012; Allender et al. 2006; Forzán and Wood 2013; Hamed et al. 2013; Sutton et al. 2014). Given the apparent correlation between disease-related mortality and infection prevalence with FV3-like ranaviruses (Haislip et al. 2011; Hoverman et al. 2011; Brenes et al. 2014b), Gray and Miller (2013) suggested that prevalence >40 % in amphibian populations might signal that an outbreak is occurring. Although these rules of thumb may be useful in interpreting prevalence levels, we urge caution in interpreting prevalence data outside of the broader biological

context. It is also worth noting that ranavirus die-offs can occur quickly (<2 weeks; Todd-Thompson 2010; Waltzek et al. 2014), so frequent sampling is necessary to detect and understand the epizootiology of ranaviruses.

Lastly, it is important to recognize that infected individuals may be more or less likely to be detected or captured than uninfected individuals, which can bias prevalence estimates (Cooch et al. 2012). For instance, moribund fish and tadpoles are often found near the surface thus can be detected easily and inflate prevalence, whereas sick turtles may move less and have lower detection probabilities resulting in underestimates of prevalence. Variation in detection probabilities through time (e.g., developmental stages) and among locations also can lead to apparent differences in prevalence that do not reflect actual differences in the proportion infected. While we are unaware of any ranavirus surveillance studies that accounted for detection probabilities, we think that doing so will substantially improve our understanding of ranavirus biology.

Infection prevalence is useful when describing the distribution of ranaviruses among regions and host species, but it does not convey information about risk or rates of infection. Infection incidence is the rate at which individuals become infected with a pathogen (i.e., the number of new cases that occur in a specified time period; Wobeser 2006). In small captive populations, it may be possible to determine how many individual animals become infected over short intervals. For example, if an initial survey found that 2 of 50 animals were infected and a second survey at the end of the month found that ten individuals were infected, then the incidence was 16.7 % (=8 new cases/48 at risk) per month. Note that individuals infected at the beginning of a study are not at risk of developing the infection so they are not included when estimating incidence. If populations are not closed (i.e., immigration or emigration occurs), calculations of incidence rates need to be adjusted for the time at risk (i.e., see Dohoo et al. 2003 for details).

Estimating incidence in wild populations is difficult, because we generally cannot track the fate or infection status of individuals. Two approaches are often used to estimate infection incidence in wild populations. First, sentinels can be used, which are uninfected individuals that are introduced into the environment (e.g., tadpoles in cages placed in a pond) and regularly screened for infection. Sentinel species should be highly susceptible to ranavirus, such as the wood frog (*Lithobates sylvaticus*) in North America (Hoverman et al. 2011). A second approach is capture-mark-recapture (CMR) studies where individuals are given unique marks and released. During subsequent encounter periods (e.g., trapping, netting), the researcher records the new and recaptured individuals, and determines their infection status. Given that individuals are released, infection status must be determined using nonlethal methods (St-Amour and Lesbarrères 2007; Gray et al. 2012). Ultimately, CMR models estimate the probability of individuals changing infection status while accounting for imperfect detection of the pathogen and imperfect recapture probability of the host.

## 2.2 *Planning Surveillance*

Cross-sectional studies that sample multiple populations during one time period are most appropriate for understanding the distribution of ranaviruses, while longitudinal studies that sample the same populations through time are most useful in understanding the epidemiology of ranaviruses and their impacts on populations. For organizations starting surveillance programs, we recommend starting with a cross-sectional study that involves widespread sampling across multiple taxa with the goal of identifying locations with elevated levels of ranavirus infection. If funding is limited, species with known high susceptibility to ranavirus or species of conservation concern can be targeted. Lethal samples (i.e., organ tissue) will likely result in greater detection of ranavirus compared to nonlethal samples (i.e., swabs, tail-clips; Gray et al. 2012). Once populations or sites with high ranavirus prevalence are identified, a more intensive longitudinal study can be performed that involves frequent sampling through the annual cycle to understand seasonal and annual trends. Sampling once every 2 weeks while hosts are present should be sufficient to detect most outbreaks (Todd-Thompson 2010).

Ranavirus outbreaks can occur because of natural or anthropogenic factors (Gray et al. 2009). Some known natural factors are host density, species composition, temperature, and host development (Gray et al. 2009; Brunner et al. 2015). Anthropogenic factors could be related to stressors (e.g., pesticides, Kerby et al. 2011) or the introduction of novel isolates (i.e., pathogen pollution, Storfer et al. 2007). Thus, to identify the causal factors for outbreaks, ideally host densities and stages of development, water and ambient temperature, and water quality should be measured during surveillance programs. If ranavirus is detected, it can be isolated from fresh or frozen tissue (Miller et al. 2015), and genomic comparisons can reveal whether it is a novel isolate that was potentially introduced (Jancovich et al. 2015b).

Understanding the impacts of ranavirus on populations is a fundamental conservation question (Duffus et al. 2015). Sampling the same sites over several years is necessary to understand possible population impacts (e.g., Price et al. 2014). In addition to sampling individuals for ranavirus infection, mark-recapture methods (e.g., Jolly-Seber) can be used to estimate host population size (Williams et al. 2002). Estimates of prevalence, incidence rate, and host abundance are essential to make informed decisions on ranavirus impacts and to identify causes of outbreaks so intervention strategies can be implemented.

## 3 *Study Design*

When designing a surveillance study, sites to be sampled should be selected randomly unless certain sites need to be targeted because they are of key conservation interest. Random sampling could be stratified based on different geographic areas or

a hypothesis related to ranavirus emergence, such as human land-use (e.g., agricultural vs. forested). Random site selection avoids unintentional biases and potential confounding factors. For instance, sites that are easily accessed will be easier to sample, but may also have greater rates of visitation by others (e.g., people fishing), which could increase rates of ranavirus introduction or levels of stressors. The number of sites sampled will depend on the study's objectives and available resources. Clearly, as number (and spatial extent) of sites increases, your conclusions will be more general. However, there also is merit in performing intensive sampling at a few sites, especially those with known reoccurring die-offs.

When sampling, individuals should be randomly collected. Ideally, captured individuals should be placed into separate numbered containers, and the individuals that are processed should be selected using a random numbers table or statistical software. Individuals should not be cohoused because transmission of ranavirus can occur rapidly between them (Brunner et al. 2007; Robert et al. 2011). Another approach is to process individuals as they are captured until a target sample size is met. Importantly, individuals that are processed should not be haphazardly selected from a group, because bias can be introduced (Gotelli and Ellison 2004). If morbid individuals are observed, they can be targeted for diagnostic purposes (Miller et al. 2015); however, targeting individuals with possible gross signs of ranaviral disease may overestimate prevalence or incidence rate. Alternatively, if the goal of surveillance is to declare a site as "ranavirus-free" (Sect. 7), targeting apparently morbid individuals can increase the probability of detecting the pathogen.

Lastly, surveillance studies in wild populations are important to learn about the distribution of ranavirus and effects on host populations. However, identifying factors responsible for outbreaks in wild populations can be challenging. Laboratory and mesocosm studies can be useful in identifying natural and anthropogenic factors that facilitate emergence. Information from controlled studies can be used to design surveillance studies that target certain hypotheses for ranavirus emergence. Additionally, controlled studies can inform field personnel of factors that should be measured (e.g., water quality) in conjunction with infection status and population abundance.

## 4 Required Sample Size

Determining the number of samples that need to be collected is generally a first step in designing a surveillance project. Required sample size will depend on whether your goal is to (1) detect the pathogen or (2) obtain a precise estimate of prevalence that can be used for statistical inferences. To estimate sample size necessary to detect a pathogen, you need (1) a previous estimate or assumed level of prevalence, (2) estimate of host population size, and (3) a specified level of confidence (generally 95 %) in detecting the pathogen (Amos 1985; Thoesen 1994). As prevalence of

**Table 1** Required sample size<sup>a</sup> to detect ranavirus in a host population with 95 % confidence given the population size and assumed infection prevalence

Estimated population size	Assumed infection prevalence of ranavirus			
	20 %	10 %	5 %	2 %
50	5	20	35	50
100	8	23	45	75
250	11	25	50	110
500	13	26	55	130
2,000	15	27	60	145
>100,000	15	30	60	150

<sup>a</sup>Calculated following methodology in Amos (1985) and Thoesen (1994)

the pathogen decreases and host population increases, the sample size required to detect a pathogen increases (Table 1). Thus, a small sample size ( $n \leq 10$ ) is required to detect an outbreak of ranavirus; however, a large sample size is required ( $n=35\text{--}150$ ) to detect ranavirus if prevalence is low ( $\leq 5\%$ ). In general, we recommend a minimum sample size of 30 per site for widespread surveillance projects that are attempting to detect ranavirus (Table 1). Larger sample sizes should be collected at sites of concern where precise estimates are needed to identify factors associated with emergence.

Determining the required sample size to obtain a precise estimate of prevalence requires: (1) a previous estimate of prevalence ( $\hat{p}$ ), (2) a specified level of error ( $d$ ) that you are willing to tolerate in the estimate of prevalence, and (3) a specified level of confidence in the prevalence estimate (generally 95 %). Sample size can be estimated as,

$$n = \hat{p}(1 - \hat{p}) \left[ \frac{1.96}{d} \right]^2,$$

(1)

where 1.96 is the critical value for the standard normal curve at 95 % confidence. If a previous estimate of  $\hat{p}$  is unavailable,  $\hat{p} = 0.5$  can be used. Thus, if  $\hat{p} = 0.85$  and  $d=0.05$ , required  $n=196$ . However, if you are willing to accept a larger error in estimating  $\hat{p}$  (e.g.,  $10\% = 0.10$ ), required  $n=49$  when  $\hat{p} = 0.85$ . Additionally, as  $p$  approaches 0.5, the required sample size for a precise estimate of  $\hat{p}$  increases. In the previous example where  $d=0.10$ , required  $n=96$  for  $\hat{p} = 0.5$ .

Detecting differences in prevalence between two sites with a statistical test can require a large sample size (Table 2). For example, required  $n=219\text{--}408$  per site to detect a 10 % difference in prevalence between two sites with 95 % confidence ( $\alpha$ ) and statistical power ( $\beta$ )=80 % depending on the value of the two proportions. Several websites (<http://epitools.ausvet.com.au/content.php?page=2Proportions>) and software packages are available for planning required sample size considering

**Table 2** Required sample size<sup>a</sup> for detecting differences between two proportions with 95 % confidence ( $\alpha=0.05$ ) and 80 % statistical power ( $\beta=0.80$ )

Proportion 1	Proportion 2										
	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
0.05		474	88	43	27	19	14	11	8	7	6
0.1	474		219	72	38	25	17	13	10	8	7
0.2	88	219		313	91	45	28	19	13	10	8
0.3	43	72	313		376	103	49	29	19	13	11
0.4	27	38	91	376		408	107	49	28	17	14
0.5	19	25	45	103	408		408	103	45	25	19
0.6	14	17	28	49	107	408		376	91	38	27
0.7	11	13	19	29	49	103	376		313	72	43
0.8	8	10	13	19	28	45	91	313		219	88
0.9	7	8	10	13	17	25	38	72	219		474
0.95	6	7	8	11	14	19	27	43	88	474	

<sup>a</sup>Sample size provided is per proportion and calculated using <http://epitools.ausvet.com.au/content.php?page=2Proportions>

the minimum detectable difference between proportions,  $\alpha$  and  $\beta$ . Required sample size decreases as the minimum detectable difference increases, and confidence level and power of a statistical test decrease.

5 Data Analysis

5.1 Confidence Intervals

Even with large sample sizes, there is uncertainty associated with any estimate of prevalence or incidence. Confidence intervals are a common measure of conveying the degree of certainty in these estimates. They can also be used to statistically compare estimates of prevalence, where nonoverlapping confidence intervals imply a statistical difference. To construct a confidence interval for incidence, the proportion is divided by the time interval.

A common approach to estimating confidence intervals is based on the standard normal approximation. This process involves calculating the standard error of a proportion, multiplying by the critical value associated with 95 % confidence for the standard normal distribution (1.96), and adding and subtracting this product from the sample estimate for prevalence (Brown et al. 2001):

$$\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{2}$$

This approximation should only be used if sample size is large ( $n>20$ ) and  $0.10 < \hat{p} < 0.90$  ; otherwise, confidence intervals can extend beyond 0 and 1, which is nonsensical (Brown et al. 2001).



There are several better methods for estimating confidence intervals for proportions (reviewed in Brown et al. 2001). We recommend the Wilson score interval (Wilson 1927), because it performs well at lower sample sizes, when  $\hat{p}$  is near 0 or 1, and it is not overly conservative (as with some continuity correction methods). The equation is:

$$\frac{1}{1 + \frac{1.96^2}{n}} \left[ \hat{p} + \frac{1.96^2}{2n} \pm 1.96 \sqrt{\frac{\hat{p}}{n} (1 - \hat{p}) + \frac{1.96^2}{4n^2}} \right], \quad (3)$$

with the same variables as (2). Hand calculation can be time consuming; however, many statistical packages estimate the Wilson score interval (e.g., the R package “binom”) and some websites are available (<http://vassarstats.net/prop1.html>). Appendix 1 provides example code in R for estimating confidence intervals.

## 5.2 Comparing Proportions

While it is useful to describe the degree of confidence in an estimate of prevalence or incidence, we are more often interested in comparing these estimates between groups or populations. Chi-square tests are often used to compare proportions among populations; the most common is the Pearson’s chi-squared test:

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}, \quad (4)$$

where  $O_i$  is the observed number of infections for population  $i$ , and  $E_i$  is the expected number of infections for population  $i$  according to the null hypothesis. Generally, the null hypothesis is that infection prevalence is equal among populations. For example, consider the scenario where 10 of 35 animals tested positive for ranavirus in one population and 20 of 45 tested positive in another. The contingency table is:

	Population A	Population B	Total
Infected	10	20	30
Not infected	25	25	50
Total	35	45	80

The expected infection prevalence, assuming no difference among populations, would be  $(10+20)/(35+45)=0.375$ . Thus, the number of infections one would expect in each population would be  $0.375 \times 35$  animals = 13.125 in the first population and  $0.375 \times 45$  animals = 16.875 in the second. The  $\chi^2$ -test statistic is the sum of the squared differences between observed and expected values divided by the

expected value for populations  $i = 1, 2, 3, \dots, n$ . This statistic is compared to a critical value from the chi-squared distribution with  $(\text{row} - 1) \times (\text{columns} - 1)$  degrees of freedom for evidence that infection prevalence is different in at least one population. Here there are two rows for infected and uninfected, and two columns for the two populations, so the degrees of freedom  $= (2 - 1) \times (2 - 1) = 1$ . If the test is significant and there are  $>2$  populations, subsequent pairwise comparisons can be performed following the same methodology, with appropriate correction of experimentwise error rate (e.g., Bonferroni correction). Chi-square tests require that no more than 20 % of expected counts are  $<5$ , which may not be achieved especially in populations with low infection prevalence. If one margin of the contingency table is fixed (e.g., if the number of samples from sites A and B in our example were set *a priori* at 35 and 45, respectively), then Barnard's exact test is a powerful alternative to the chi-square test that avoids the problem with low expected counts (Martín Andrés et al. 2004), and can be performed using the "Barnard" package in R. Appendix 1 provides example code in R for testing for differences in proportions.

Logistic regression is a robust and more flexible framework for comparing the probability of infection (or death) among individuals or populations given environmental or host characteristics. The logistic model predicts the logit-transformed probability of a binary outcome (e.g., infection, mortality) as a linear function of one or more predictor variables:

$$\text{logit}(p_i) = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i}, \quad (5)$$

where  $\beta_0, \dots, \beta_m$  are the intercept and regression coefficients for predictor variables  $x_1, \dots, x_m$ . The logit transform is the log of the odds ratio, where the odds ratio is calculated by exponentiating both sides of (5).

$$\exp(\text{logit}(p_i)) = \left(\frac{p_i}{1-p_i}\right) = e^{(\beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i})} \quad (6)$$

If  $x_1$  is a categorical predictor (e.g., male vs. female), then  $\exp(\beta_1)$  can be interpreted as the increased (or decreased) odds of infection for males relative to females. If  $x_1$  is instead a continuous variable (e.g., animal length), then  $\exp(\beta_1)$  is the increased (or decreased) odds of infection with a one-unit increase in the predictor variable. It is important to be careful when interpreting odds-ratio coefficients relative to the units measured (e.g., mm vs. cm) as well as in the context of the range of values that were measured. For example, a large predicted increase in risk with each centimeter may seem impressive, but if all of the animals measured were within 0.1 cm of each other, the actual effect size is much less substantial.

Logistic regression can also be used to estimate the risk factors associated with ranavirus occurrence among populations (e.g., Gahl and Calhoun 2008; Greer and Collins 2008). For instance, you may be interested in finding the predicted probability of ranavirus infection or a die-off occurring in particular populations. It is possible to

use the coefficients of the logistic regression to predict this probability for population  $i$  (or analogously, individual  $i$ ) as:

$$P_i = \frac{e^{\beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i}}}{1 + e^{\beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i}}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1,i} + \dots + \beta_m x_{m,i})}}. \quad (7)$$

Suppose we fit a logistic regression model predicting the occurrence of ranavirus in wetlands as a function of distance from the nearest road where the intercept was  $\beta_0 = -0.5$  and the slope parameter for distance was  $\beta_1 = -0.1$ . In this case, a pond that was 10 km from a road would have a predicted probability of ranavirus occurrence as:  $1/(1 + \exp[-(-0.5 + -0.1 \times 10)]) = 0.182$ , while a population that was 5 km away would have a predicted probability of  $1/(1 + \exp[-(-0.5 + -0.1 \times 5)]) = 0.269$ . Most statistical packages will provide predicted probabilities and confidence intervals from a logistic regression model. Appendix 1 provides example code in R for logistic regression.

### 5.3 Viral Titers

The above statistical approaches classify infection as binary; an individual is infected or not. However, infection can be thought of as a continuum from subclinical to clinical infections, where the latter is resulting in disease and possibly mortality (Miller et al. 2015). Quantitative PCR and cell culture-based methods (e.g., plaque assays and TCID<sub>50</sub>) are common techniques to estimate viral titers in tissue (Miller et al. 2015). Inasmuch as viral titers in tissues correlate with the severity of infection, these data provide additional insight into the possible effects of ranavirus on populations. Consider, for instance, measuring ranavirus prevalence and titers through time in a population of a tolerant species (e.g., American bullfrog, *Lithobates catesbeianus*; Hoverman et al. 2011). One might observe that the prevalence of ranavirus was quite high, but titers were very low. If changing conditions (e.g., rising temperatures) were hypothesized to make this species more susceptible, then one would expect to see viral titers increase with increasing temperature, while prevalence of infection would remain unchanged.

Vital titers are often reported as log<sub>10</sub>-transformed values of virus concentration per unit of genomic DNA or tissue. Such transformed titers are generally normally distributed, which are suitable for simple linear models (i.e., regression, analysis of variance). For example, the relationship discussed above could be tested with a linear regression of viral titers on temperature.

Because the log of zero is undefined, it is common practice to add one or a number representing the minimum detectable level to all numbers including zero before taking the log. If there are many zeros (i.e., individuals that tested negative) in the dataset, the resulting distribution will not be normal. However, if you are interested in the distribution of titers in infected animals only, it is appropriate to exclude the

zeros for the uninfected individuals. Alternatively, you can use zero-inflated models. These models account for the probability that an individual is infected using the equivalent of a logistic model, and given that an animal is infected, predicts the number of virions with typically a Poisson or negative binomial distribution. These models can also be applied to other surveillance data, such as the number of infected animals in a population, where there may be zeros because there is no infection in the population or because infected animals were missed during sampling. We direct the readers to Dohoo et al. (2003) and Zuur et al. (2012) for additional guidance.

## 5.4 Analysis of Survival Data

While we are often interested in the probability of infection or death, it is also useful to understand the timing or rate of mortality. In survival analyses, the fate of specific individuals is followed over time at frequent intervals; thus, these designs are probably most appropriate in captive populations (e.g., zoos, laboratory studies), where every individual can be checked regularly. When the fate of all individuals is known over time, survival can be represented as a curve ranging from 100 to 0 % over the duration of the study.

Censoring is when the fate of some individuals during a study is unknown, and must be accounted for in survival analyses. Right censoring occurs when the fate of an individual is not observed after some point in time; the individual is censored after its last observation. Right censoring also occurs when individuals are euthanized during or at the end of a study to collect diagnostic information. If an animal was infected at some unknown time before the start of the study, it is left censored. In field studies, it is common that individuals are added to a study after the first sampling date, which is called staggered entry.

Information on the fate of individuals at risk at each time point (i.e., excluding those that have been censored) is used to estimate time-specific survival,  $S(t)$ , and analyzed with various statistical packages (e.g., Program MARK, <http://www.phidot.org/software/mark/>). One of the most common survival estimators is the Kaplan–Meier (K–M) function:

$$S(t) = \prod_{t_i < t} \frac{n_i - d_i}{n_i} \quad (8)$$

where  $S(t)$  is probability of surviving until time  $t = t_i$ ,  $n_i$  is the number of individuals that survived and were not censored before time  $i$ , and  $d_i$  is the number that died at time  $t_i$  (see Jager et al. 2008 for an overview). The probability of surviving up to time  $t$  is the product of the current and previous survival probabilities.

The K–M survival estimates can be compared between two groups of samples using the Mantel–Haenszel test, which is essentially a contingency table approach

(Sect. 5.2), where expectations and deviations are calculated through time. The contingency table is:

	Group A	Group B	Total
Event	$d_{Ai}$	$d_{Bi}$	$d_i$
No Event	$n_{Ai} - d_{Ai}$	$n_{Bi} - d_{Bi}$	$n_i - d_i$
At Risk	$n_{Ai}$	$n_{Bi}$	$n_i$

where  $i$  refers to the time  $t_i$  and subscripts A and B are the two groups. The *expected* number of deaths at time  $t_i$  in group A, if both groups are identical in terms of their survival functions, is:

$$\hat{e}_{Ai} = (d_i \times n_{Ai}) / n_i. \quad (9)$$

Now, the expected number of deaths can be compared to the *actual* number of deaths in group A at time  $t_i$ , and repeated over  $i = 1, 2, 3, \dots, m$  sample periods. In the case of comparing two groups, expectations can be calculated for one group, because deviations in group A imply deviations in group B. The test statistic is:

$$Q = \frac{\left( \sum_{i=1}^m d_{Ai} - \sum_{i=1}^m \hat{e}_{Ai} \right)^2}{\sum_{i=1}^m \hat{V}(\hat{e}_{Ai})}, \quad (10)$$

where  $V$  is the variance of the expected number of deaths. The test statistic is chi-square distributed with one degree of freedom (Dohoo et al. 2003; Hosmer et al. 2008).

While the Mantel–Haenszel test is relatively easy to calculate, it cannot accommodate more than two groups or continuous predictors or covariates. The Cox Proportional Hazard (Cox PH) model is a more general method of testing for differences in survival curves among groups, or among individuals with continuous covariates (e.g., body size). Cox PH estimates a baseline *hazard* function (Box 1), and tests whether individuals in the groups have a higher or lower hazard than the baseline (Hosmer et al. 2008). Cox PH nonparametrically estimates a baseline hazard function,  $h_0(t)$ , from the data. The hazard for an individual with covariates  $x_1, x_2, \dots, x_n$  is:

$$h_0(t) e^{[\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n]}. \quad (11)$$

When the linear portion of the model in brackets is equal to zero, the exponential term is one and the hazard is equal to  $h_0(t)$ , the baseline hazard. If the sum of the terms in brackets is  $>0$ , then the hazard increases by some proportion; if it is  $<0$ , the hazard is reduced by some proportion. For example, if the coefficient for

females (relative to males) was  $\beta_{\text{Female}}=0.693$ , then females would have a hazard that was  $\exp(0.693)=2\times$  greater than that of males. In Cox PH analyses, the focus tends to be on the proportional differences in survival between groups, although it is possible to extract the baseline hazard from most statistical packages.

There are limitations to the Cox PH model. First, it cannot accommodate left-censored observations. Second, it assumes the proportional difference in hazard between groups (e.g., males vs. females) is constant through time. Thus, if survival curves are plotted, they should not cross or diverge; they should be approximately parallel through time. If meeting either of these assumptions is unreasonable, readers should consult a text on survival analyses (e.g., Dohoo et al. 2003; Hosmer et al. 2008) or statistician for alternative approaches.

One alternative approach to Cox PH is accelerated failure time (AFT) models, sometimes called parametric survival models (Hosmer et al. 2008). There are two key differences between AFT and Cox PH models. First, in AFT models, the functional form (but not rates) of the underlying hazard is specified *a priori* rather than estimated from the data (Box 1). For instance, a constant hazard would be modeled using an exponential model (Hosmer et al. 2008). Because the form of the hazard is set *a priori* and only the model parameters are estimated, survival estimates can be predicted beyond the observed time period and may have more statistical power.

## 5.5 Mark-Recapture Studies

Many of the difficulties inherent in estimating epidemiologically relevant parameters in wildlife populations (e.g., individual fates, population size) can be addressed using CMR methods (reviewed in Cooch et al. 2012). This is an active area of research

### Box 1

**Hazard function,  $h(t)$** —instantaneous rate of death at time  $t$ . The cumulative hazard is written as  $H(t)$ .

**Survival function,  $S(t)$** —probability of surviving beyond time  $t$ .

**Probability density function,  $f(t)$** —the expected distribution of times to death.

These functions are related to each other:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{\partial \ln S(t)}{\partial t}$$

$$f(t) = S(t)h(t)$$

$$S(t) = \exp\left[-\int_0^t h(t) dt\right] = \exp[-H(t)].$$

and one with a large literature (e.g., Amstrup et al. 2005; Thomson et al. 2009). Thus, we will simply provide an overview of approaches that may prove useful to understanding ranavirus epidemiology and direct the reader to the literature above.

Closed population models are particularly useful for estimating population size (or density) and prevalence of infection. These models assume that the initial and subsequent recapture sessions occur close enough in time that one can assume there has been no birth, death, immigration, or emigration. In the simple case where there are two capture occasions, the population size,  $\hat{N}$ , is estimated by the actual count of individuals,  $C$ , adjusted for the detection probability,  $\hat{p}$  (i.e., the Lincoln–Peterson estimator):

$$\hat{N} = C / \hat{p}. \quad (12)$$

The detection probability is estimated as the fraction of initially marked individuals that are recaptured. This model can be extended to account for multiple capture sessions as well as differences between groups (e.g., males vs. females) or states (e.g., infected vs. uninfected). Importantly, the detection probability can be modeled separately for different groups or states, which allows you to account for differences in detection probabilities between ranavirus-infected and -uninfected or symptomatic and asymptomatic animals (see Sect. 2.1).

Open CMR models do not assume that the population is closed to demographic changes and are generally better suited for repeated monitoring and estimating demographic parameters, particularly apparent survival,  $S$ . Parameters in open CMR models can be modeled separately between groups (e.g., infected and uninfected) or as a function of covariates (e.g., age, size), which provides a means of estimating the impact of disease on individuals in natural settings. One could, for instance, determine whether apparent survival differs between ranavirus-infected and -uninfected fish, and whether these differences are constant between adults and juveniles. In a similar framework, it may be possible to estimate the population growth rate as a function of the occurrence or prevalence of disease (Cooch et al. 2012).

Multi-state models are an extension of CMR models that allow individuals to transition between different states (e.g., uninfected and infected). This powerful modeling approach provides a means of estimating the rate or probability of transitioning from uninfected to infected states (i.e., incidence) and vice versa. These models assume that survival and transitions between states are temporally separated (e.g., individuals first survive then become infected). Additionally, only one transition (e.g., uninfected to infected) can occur between encounter events. Thus, careful design of a CMR study is essential. These and related models can be extended to account for misclassification of states (e.g., infection status is not measured perfectly) or partial observability (e.g., the individual is observed but its infection status is not determined). Considering the complexity of working with CMR models, we recommend consulting a statistician during study design and analysis.

## 6 Use of Dynamic Models

Dynamic models can be very useful in studying host–pathogen interactions. Within-host models can elucidate physiological mechanisms that lead to host infection and disease (e.g., Mideo et al. 2008, 2011; Woodhams et al. 2008). In comparison, between-host models focus on the fate of individuals and populations when a pathogen is introduced or circulating (Hastings 1997). In this section, we will focus on the latter because of their usefulness in predicting the effects of pathogens on populations. To date, few dynamic models have been formulated for ranaviruses (e.g., Duffus 2009). Thus, several of our examples will come from the wildlife disease literature and modeling efforts with the emerging pathogen *Batrachochytrium dendrobadiitis* (Bd).

### 6.1 SI/SIR Models: Transmission

Susceptible-infected-recovered (SIR) models examine transmission dynamics using a series of ordinary differential equations that model and predict one of three outcomes: pathogen extinction, host extinction, or pathogen–host persistence (Allen 2006). In many simple cases, the total population of hosts is divided into three subpopulations: individuals susceptible to infection ( $S$ ), infected individuals ( $I$ ), and individuals that have recovered ( $R$ ) from infection and cannot be re-infected or at least have temporary immunity.  $R$  can also be the individuals removed from the population. A simpler version of the model is where individuals cannot become immune, the susceptible-infected (SI) model (Allen 2006). In this version, if individuals clear the infection, they become susceptible again. Here, we describe the basic SIR model.

In the simplest continuous time SIR model, the total population size ( $N$ ) can be assumed constant

$$N = S + I + R \quad (13)$$

where  $S$ ,  $I$ , and  $R$  represent the number of individuals in each respective subpopulation (Hastings 1997). The rate of change of each subpopulation at time  $t$  can be modeled as

$$dS / dt = -\beta SI \quad (14)$$

$$dI / dt = \beta SI - \gamma I \quad (15)$$

$$dR / dt = \gamma I \quad (16)$$

where  $\beta$  is the rate at which hosts contact and transmit the infection to each other and  $\gamma$  is the host recovery rate (or removal rate). Here, transmission is assumed to



be density-dependent, as transmission is represented as  $\beta SI$ . Some evidence exists that transmission of ranavirus may be density-independent (Harp and Petranka 2006), and can be modeled as  $\beta I/N$ . McCallum et al. (2001) provide other forms of transmission functions, including density-independent transmission and nonlinear functions of density. Because demography (birth, death, immigration, or emigration) is not included in this model, the only equilibrium occurs when all individuals are in the susceptible class (with  $I=0$ ). For an epidemic to occur, the number of infected individuals must increase  $dI/dt > 0$ . The reproductive number of a disease ( $R_0$ ) is the number of secondary cases that one infected individual would produce on average in a susceptible population, and is equal to

$$R_0 = \beta S / \gamma. \quad (17)$$

If  $R_0 > 1$ , number of infections are increasing in a population, and is representative of an epidemic. However, due to the density-dependent nature of this model, there is a minimum population size for an epidemic to occur (the threshold population size is  $N_T = \gamma/\beta$ ), and the epidemic ends before all susceptible individuals become infected (Hastings 1997). When modeling epidemics, the time scale is assumed short enough to ignore births and other forms of mortality in the host population. This assumption can be relaxed in more complex models by adding births to the susceptible population and natural mortality to each subset of the population.

For ranavirus and most natural populations, the basic SIR model is likely too simplistic. Duffus (2009) used a discrete-time SI model to show that ranavirus could be maintained in a population of common frogs (*Rana temporaria*) in the UK with only transmission between adults. Her model included natural and disease induced mortality and recruitment from earlier life stages. The transmission rate was determined by the contact rate between adults and the likelihood of being infected given contact. Duffus (2009) also demonstrated that transmission between adults could maintain two syndromes of ranavirus (the ulcerative and hemorrhagic forms) in a single population. These models showed the conditions that could result in persistence of ranavirus in populations of common frogs, and which parameter estimates need additional data to better understand the system and predict outcomes in particular populations (Duffus 2009). Another model is in development for wood frogs (*Lithobates sylvaticus*) that investigates stage-specific susceptibility and waterborne transmission to recreate die-off patterns observed in natural populations (JLB, unpublished data).

Other model expansions could be particularly useful for predicting ranavirus dynamics in natural populations. For example, most ranavirus host species exist in communities where they are likely to interact with other susceptible species, possibly from different ectothermic vertebrate classes (Gray et al. 2009). Brenes et al. (2014a) demonstrated that interclass transmission of ranavirus through water was possible. He also showed that ranaviral disease outcomes depended on species composition in the amphibian community and which species was initially infected with ranavirus (Brenes 2013). These studies could serve as a starting point for determining transmission probabilities in aquatic communities with multiple species.

In other disease systems, the addition of multiple species to transmission models had an effect on the focal host population, but depended on the host's competency as a reservoir and its dominance within the community (Keesing et al. 2006). The addition of multiple host species can make the analysis of SIR models challenging. To date, most models have included only two species and the pathogen (Keesing et al. 2006), which may be unrealistic for some ranavirus–host systems. Dobson (2004) dealt with the large number of parameters in multi-species models by scaling the parameters as allometric functions of host body size, although it is unclear such a relationship exists with transmission of ranavirus. L  lu et al. (2013) provide an example of a model including trophic transfer of a parasite (*Toxoplasma gondii*) from rats to cats and vertical transmission in cats. Similar complex interactions certainly occur among ranavirus hosts species, such as predation or necrophagy, and mechanical transmission by mosquitoes has been hypothesized (Allender et al. 2006; Johnson et al. 2007; Kimble et al. 2014). Despite the large number of possible interactions in a ranavirus–host system, several interactions are likely unimportant to its epidemiology. One strategy would be to create several competing models and fit them to data on dynamics in natural populations or in mesocosm studies to identify the most important mechanisms for transmission.

For researchers interested in using SIR models to examine ranavirus, we recommend Otto and Day's (2007) book *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*, which reviews the mathematics and describes the process necessary for constructing and analyzing models primarily with ordinary differential equations. An understanding of computer programming and use of software (e.g., Matlab, Maple, Mathematica, R) will be necessary to construct models and perform simulations for most analyses. Appendix 2 provides example code in Matlab for a simple SIR model.

## 6.2 Individual-Based Models/Pattern-Oriented Modeling

Individuals-based models (IBMs), sometimes called agent-based models (ABMs), are also very useful for examining disease dynamics. IBMs are simulation-based, and during each time step, a set of rules or probabilistic events occurs involving each individual. IBMs are often easier for biologists to construct than SIR models, because they do not require solving differential equations. However, IBMs can be complex and require computer programming skills. These models often operate on a set schedule of events that are implemented using sequential equations, a series of for-loops, and if-then statements that determine an individual's actions or fate. For disease IBMs, each individual's disease state is recorded and their risk of infection can depend on their interaction with other individuals or the environment. There are also other types of IBMs that use differential equations. For example, Briggs et al. (2010) developed an IBM with differential equations that explicitly incorporated individual Bd load and further examined how a pathogen reservoir and a long-lived tadpole stage affected whether the frog population could persist with Bd

or experience local extinction. Similar models could be developed for ranavirus that include viral load and shedding to better understand how the virus might interact with the host and factors that initiate die-offs. One attractive aspect of IBMs is that they can explicitly incorporate animal behavior. For ranavirus, researchers might be interested in how different behaviors, such as schooling or necrophagy, affect host populations and persistence with the pathogen.

A useful technique for creating IBMs and determining plausible interactions is called pattern-oriented modeling (POM). In POM, data are used to determine several salient patterns seen in a natural system of interest that form the basis of model evaluation. Multiple possible forms of an IBM are created, representing different hypotheses about host–pathogen interactions. The different IBMs are evaluated based on their ability to recreate the salient patterns (Grimm et al. 2005; Grimm and Railsback 2012). When a model is able to match multiple patterns, it is more likely to be structurally realistic (Wiegand et al. 2003), and capable of producing testable predictions. In using POM, researchers can also contrast different hypotheses, determine a useful model structure, and reduce parameter uncertainty.

For researchers interested in developing IBMs, we recommend two books: Grimm and Railsback’s (2005) *Individual Based Modeling and Ecology* and Railsback and Grimm’s (2011) *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Both titles describe a “best model practice” called object-oriented design and description (ODD), which is a standard format to describe various aspects of an IBM. The latter title goes through the process of building IBMs with examples and code for a relatively user-friendly and free program called NetLogo (<http://ccl.northwestern.edu/netlogo/index.shtml>). NetLogo includes a library of preconstructed models, including AIDS, Disease Solo, and Virus, which could form the basis for the development of models for ranavirus. Further, NetLogo’s website includes a Modeling Commons, where NetLogo users can share their models to help others in their own model development. Other software, such as Matlab and R, can also be used to develop and analyze IBMs.

### 6.3 Population Matrix Models

Population matrix models examine changes in population size and age structure over time. These models include parameters for the transition probability between each age class. To incorporate disease, the survival following exposure to ranavirus can be incorporated for each age class. Earl and Gray (2014) developed a stage-structured matrix model to predict the effects of ranavirus exposure during the egg, hatchling, larval, and metamorph stages on a closed population of wood frogs. This study combined information from a wood frog population model (Harper et al. 2008) with experimental challenge data (Haislip et al. 2011) to predict population outcomes. Appendix 2 provides example code in Matlab for a matrix model following Earl and Gray (2014).

Population matrix models can also be combined with transmission models to more realistically model both dynamics simultaneously. For example, Briggs et al. (2005)

merged a population model of yellow-legged frogs (*Rana muscosa*) and a SIR model of the infection dynamics of Bd based on the current knowledge of transmission and mortality rates. This model combined discrete-time between-year population dynamics with a continuous time transmission dynamics within each year. By running the model with different parameter values, Briggs et al. (2005) were able to determine which conditions resulted in extinction of the frog population, nonpersistence of the pathogen, and persistence of the frog population and the pathogen.

Population models can also be scaled up to take into account metapopulation processes. A metapopulation is a set of spatially structured local populations that periodically interact via dispersal (Marsh and Trenham 2001; Smith and Green 2005). Several ranavirus host species are likely structured as metapopulations. Metapopulation models incorporate parameters for dispersal probability between local populations as well as demographic parameters in each local population. Metapopulation models are useful to understand the spatial spread of pathogens among populations and examine the effectiveness of disease intervention strategies (Hess 1996). In amphibians, the occurrence of ranavirus outbreaks has been attributed partly to subclinically infected juveniles or adults returning to breeding sites, shedding the virus, and infecting larvae (Brunner et al. 2004). For individuals interested in population matrix models, we recommend Caswell's (2000) *Matrix Population Models: Construction, Analysis, and Interpretation*. Hanski's (1999) *Metapopulation Ecology* will be useful for those interested in investigating ranavirus effects on metapopulation dynamics.

## 6.4 Modeling Disease Intervention Strategies

One goal of modeling host–pathogen dynamics is to identify intervention strategies that thwart disease outbreaks. Currently, there are few proposed control options for ranavirus, but vaccine development is possible in the future (Miller et al. 2011). Other options include quarantining individuals or populations, culling, and creating captive populations for reintroduction if disease is likely to cause extremely high mortality to populations of conservation concern. Models also can be used to identify vulnerable points in the host–pathogen cycle that can be interrupted with intervention strategies. For example, if outbreaks are a consequence of density, emergent vegetation in wetlands can reduce the probability of transmission among amphibian larvae (Greer and Collins 2008). If stressors in the aquatic environment (e.g., high nitrogen levels) are resulting in reoccurring outbreaks, strategies that improve water quality can be used. A thorough understanding of the factors responsible for outbreaks and the ranavirus–host system is essential to identifying plausible intervention strategies. In some cases, possible intervention strategies might be infeasible to implement, excessively costly, or undesirable in natural populations. However, if strategies are feasible, models can be used to determine when and how often the strategy should be employed for the best results. SIR models and their variants can be used to explore vaccination strategies (Hethcote 2000) and other control techniques such as culling

(Lloyd-Smith et al. 2005). Cost of disease control can be incorporated into models to determine the best strategies given financial constraints (Fenichel et al. 2010). Woodhams et al. (2011) discussed possible intervention strategies for Bd and presented model results of their efficacy on individuals with and without an adaptive immunity. They also went on to show that reducing the host population size (i.e., decreasing transmission probability) could prevent extinction. For researchers interested in implementing optimal control models, we recommend Lenhart and Workman's (2007) *Optimal Control Applied to Biological Models*, which focuses on control of continuous ordinary differential equation models and includes sample code for the computer program Matlab. Optimal control can also be applied to IBMs, but effective techniques are still being developed (Federico et al. 2013).

## 6.5 Model Parameterization and "Evaluation"

There are a number of ways to parameterize models and integrate them with data. Frequently, modelers choose parameter values by searching the literature, but often not all parameter values are available. Another method is to construct a model and fit the output to an existing data sequence. In the case of ranavirus modeling, predictions could be fit to surveillance data that include abundances of infected and uninfected individuals, or the magnitude and timing of a die-off. After the model is fit to the data, the parameter values that give the best fit or that match multiple patterns (as in POM) are then used. If some parameters are known and researchers have a good idea of the possible range of other parameters, these ranges of values can be explored to determine how they change the model output. Assessing the effects of changes in parameter values is called sensitivity analysis (Cariboni et al. 2007). If the model is especially sensitive to a certain parameter, it suggests that better parameter estimation would be a valuable research direction (Biek et al. 2002; Cariboni et al. 2007), especially if the parameter estimate is not based on robust data (e.g., low sample sizes). Cariboni et al. (2007) suggest best practices for sensitivity analysis. An excellent review of parameter estimation for disease modeling of natural populations can be found in Cooch et al. (2012).

The aim of model evaluation is to determine if models typify natural systems well enough to represent the intended dynamics. This often involves determining whether or not they can be used to make accurate predictions. Frequently, the terms model evaluation, model validation, and model testing are used interchangeably. Because models are built on assumptions and simplifications, they are never truly "valid" or "correct." Augusiak et al. (2014) have suggested the term "evaluation" to represent the process of assessing the model's quality and reliability, and included six elements for proper "evaluation" of a model:

- Assessing the quality of the data used to build the model
- Evaluating the simplifying assumptions structuring the model
- Verifying that the model is correctly implemented

- Verifying that the output matches the data used to design the model
- Exploring model sensitivity to changes in parameter values, and
- Assessing whether the model can fit an independent data set not used in original model formulation.

It is recommended that model formation and “evaluation” follow a documentation procedure called TRACE (TRANSPARENT and Comprehensive Ecological documentation) that is designed to ensure reliability of models and link the science to application (Grimm et al. 2014).

## 7 Risk Analysis for Introduction of Ranavirus into an Uninfected Area

Import risk analysis (IRA) is a procedure that can be used to determine the threat of a pathogen entering a system. The consequences of pathogen introduction can be monitored directly (Sect. 2) or simulated using models (Sect. 6). The guidelines for IRA have been primarily developed from a trade perspective between two countries or regions to assess the disease risk associated with the import of live terrestrial production animals. However, the same principles can be applied to assess the risk of ranavirus introduction in wild or captive populations. In general, IRAs focus on possible infection of one species or several species within the same taxonomic class. As discussed in Duffus et al. (2015), ranaviruses are multi-species pathogens that have the capability of infecting three vertebrate classes, which makes IRA for ranaviruses complex. IRAs can be used to establish or revise trade or translocation guidelines for wildlife that could be subclinically infected with a pathogen (Smith et al. 2009). The World Organization for Animal Health (OIE) lists ranaviruses that infect amphibians as notifiable pathogens, meaning that a subsample of amphibians that are involved in international trade should be verified ranavirus negative prior to shipment (Schloegel et al. 2010). Currently, these regulations are not being enforced in most countries (Kolby et al. 2014). The procedures we outline below are based on principles and recommendations of the OIE (Vose 2000; OIE 2014), with examples of how they can be applied to parts of an IRA for the introduction of a ranavirus into an uninfected area.

### 7.1 *Defining the Hazard*

The first step in an IRA is defining an area of interest. The area could be a population of interest, such as one that contains an uncommon species that is susceptible to ranavirus, or it could be a geographic region or country (Rödger et al. 2009; OIE 2014). Generally, areas are defined based on artificial or natural barriers to animal movement or pathogen translocation (OIE 2014). For example, ranavirus virions can flow downstream in tributaries, and associated floodplains are often corridors

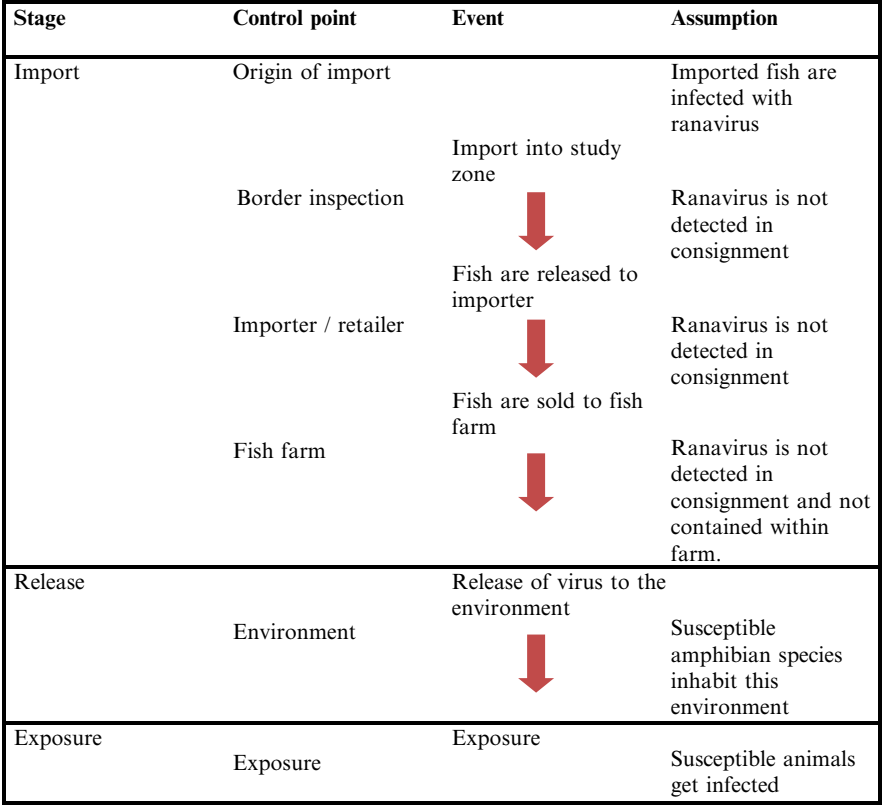
for animal movement; thus, areas should be defined by watershed for lotic systems. In lentic systems, depressional wetlands or lakes containing possible ranavirus hosts could be defined as the area of interest if it is hydrologically closed and surrounded by a terrestrial landscape. In zoological settings, the area of interest typically is the captive facility (OIE 2014).

The next step is determining the presence of ranavirus in the area of interest. Section 2 discussed surveillance studies, and additional guidelines are provided by OIE (2014). Minimum sample size to detect ranavirus depends on several factors (Sect. 2, Table 1). Additionally, infrequent sampling can result in lack of detection. Todd-Thompson (2010) showed that ranavirus in Gourley Pond of the Great Smoky Mountains National Park appeared nonexistent except for a 3-week period in late spring when an outbreak occurred resulting in widespread mortality across multiple species. Thus, sampling sites every 2 weeks when hosts are present with a large sample size ( $n > 30$ ) should result in a high detection probability. If resources are limiting, sampling at least four periods per year while hosts are present may be sufficient. Using this sampling frequency, Hoverman et al. (2012) detected ranavirus at 33 of 40 sites. Given that ranavirus could have been present at all sites in this study, a ballpark estimate of detection probability was 82.5–100 % with their sampling frequency. Sampling should be performed over several years to verify that a site is ranavirus negative. For large areas of interest, multiple sites spaced no less than the average dispersal distance of hosts should be sampled, which for amphibians is about 1 km (Wells 2007). Thus, distinct populations should be sampled without leaving large gaps between them. If ranavirus is detected, there is no reason to conduct an IRA, unless there is concern of a foreign strain of ranavirus being introduced.

Although the primary interest in the introduction of ranavirus to an area typically is for a certain species of concern, it is important that all ranavirus hosts are considered in an IRA. As discussed in Brunner et al. (2015), some hosts function as reservoirs for the virus and maintain subclinical infections resulting in low population prevalence, while other species serve as amplification hosts and initiate outbreaks. If funds are limited, a viable strategy would be to test amplification hosts, because these species tend to have lower resistance to ranavirus, and detection probabilities are therefore greater. Duffus et al. (2015) provide a list of known ranavirus hosts, and several challenge studies (e.g., Hoverman et al. 2011; Brenes et al. 2014b) can provide insight into relative difference in susceptibility between species.

## 7.2 Risk Assessment

Risk assessment involves three primary steps: identifying routes of introduction, identifying the consequence of introduction, and estimating risk. It is often useful to develop flow diagrams that illustrate each step of assessment (Figs. 1 and 2). To describe this process, below we provide an example of assessing risk to wild amphibians via import of aquacultured fish that are infected with ranavirus.



**Fig. 1** Flow diagram for possible routes of transmission of ranavirus into a naïve susceptible population of amphibians in the wild

**7.2.1 Routes of Introduction**

Routes of introduction could include dispersal paths of hosts or translocation of the virus on fomites attached to non-hosts (i.e., birds and mammals, Gray et al. 2009). Humans can play a large role in the possible introduction of ranavirus by moving between contaminated and uncontaminated sites. The environmental persistence of ranavirus in unsterile water and soil is probably at least one week (Nazir et al. 2012). Thus, recreationists that move among watersheds without decontaminating footwear or gear could be a major source of ranavirus introduction (Gray et al. 2009). Fish hatcheries are known sites of ranavirus outbreaks (Waltzek et al. 2014); thus, the release of clinically or subclinically infected fish or their effluent from the hatchery could be another major source of ranavirus introduction. For a particular area of interest, it is important to identify the most likely routes of introduction. It can be useful



to divide routes of introduction into three stages: import, release, and exposure. In the case of imported aquacultured fish, the following steps define the import stage:

- Imported fish from an infected zone are infected with ranavirus
- The infection passes undetected through border control
- The infected fish are released to the retailer
- The infected fish are sold to an aquaculture facility in the study zone.

Assuming that fish are contained in aquaculture ponds, ranavirus could be released into adjacent aquatic environments via several pathways:

- Virus contaminated effluent is released
- Infected fish escape
- Avian or mammalian predators could transport live or dead fish
- Ranavirus hosts, such as amphibians or reptiles, could enter the pond, become infected, and disperse
- Mechanical vectors, such as pets or humans, could transport the virus on fomites.

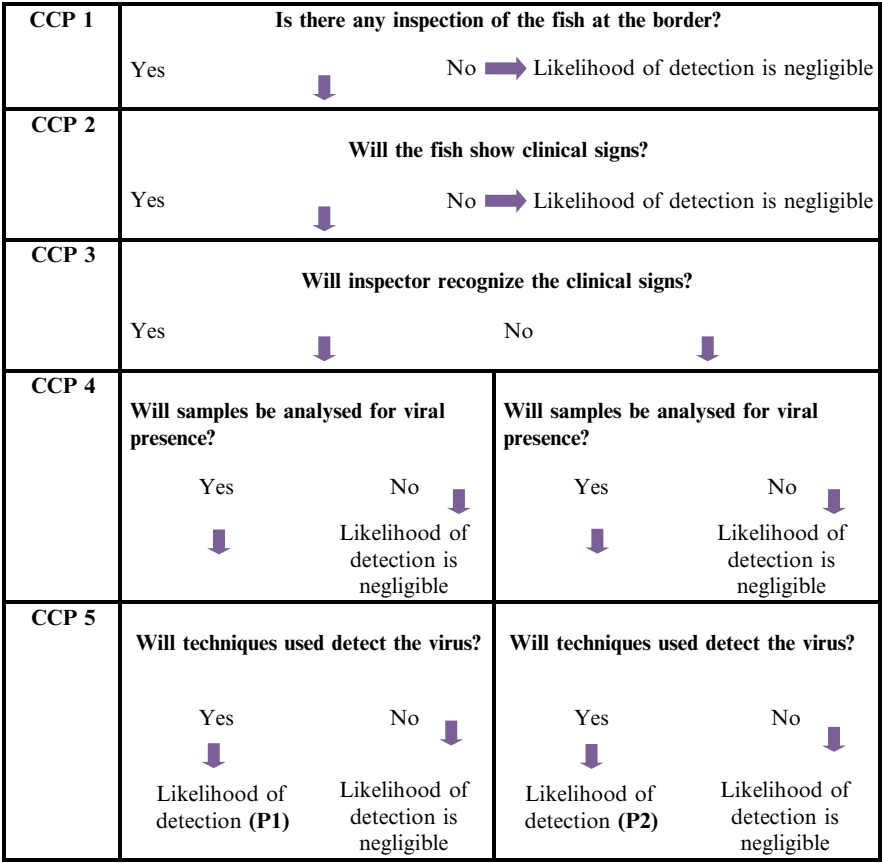
Finally, exposure to the virus could occur via several direct and indirect routes (Gray et al. 2009). Host species could be exposed to the virus in water, which is an efficient transmission medium, or the virus could be transmitted by direct contact or consumption of infected hosts (Miller et al. 2011). There is some evidence that ranavirus transmission can be density independent, which can increase extinction probabilities (Brunner et al. 2015).

### 7.2.2 Consequence Assessment

The outcome of ranavirus infection in a species can be described qualitatively or quantitatively in terms of direct or indirect consequences. Direct consequences are the effect that ranavirus has on the species of interest, which typically includes estimating the likelihood of population declines and extinction (Sect. 6). Highly susceptible species that are rare have the greatest probability of extinction (Earl and Gray 2014), especially if these species co-occur with other ranavirus hosts. Indirect consequences are costs associated with pathogen surveillance (i.e., field and diagnostic expenses) and possible repatriation of populations following extinction.

### 7.2.3 Risk Estimation

The assumption is that the virus will travel along the routes that were identified from an infected animal to a susceptible animal. In cases where it is determined that the consequence of ranavirus introduction is unacceptable, a series of critical control points (CCPs) should be established along the routes of introduction identified above, where the virus could be intercepted and the transmission terminated. The probability of the infection passing unnoticed through a CCP is estimated for each CCP by addressing several questions. This process can be summarized in a scenario



**Fig. 2** Scenario tree for detection of ranavirus in a consignment of infected fish at the border inspection. Critical control points (CCP) 1–5 are opportunities identified where the virus could be detected and future transmission terminated. P1 is the product of the “Yes” answer probabilities in the left branch of the tree. P2 is the product of the two “yes” answer probabilities in the right branch of the scenario tree. The probability of ranavirus not being detected at the border is  $1 - (P1 + P2)$

tree, where each CCP has a “yes” and a “no” branch, and a likelihood of detection is assigned (Fig. 2). In Fig. 2, CCP 1 and 4 are predetermined for each border control post, while CCP 3 will depend on the training and experience of the inspectors. CCP 2 can be affected by viral load, water temperature, and animal health. Detecting a pathogen in a laboratory test in CCP 5 is a function of two processes: sample size (Sect. 2) and performance of molecular tests (i.e., the sensitivity and specificity of PCR, Miller et al. 2015). The sensitivity and specificity of PCR for ranavirus is an ongoing research direction (Miller et al. 2015), and can be affected by sample type (i.e., lethal vs. nonlethal collection, Gray et al. 2012). In general, it is believed that liver and kidney tissue provide the most reliable estimate of detection followed by tail, toe clips, and blood (Miller et al. 2015). Assuming perfect sensitivity and specificity of PCR, the probability of detecting ranavirus is approximately 95 %

using the required sample sizes in Table 1. Risk of not detecting ranavirus in an imported consignment is calculated as:  $1 - \text{the product of the detection probabilities at all CCPs}$  (Fig. 2).

### 7.3 *Risk Management and Communication*

To manage the risk of ranavirus introduction, it is useful to perform a risk-consequence assessment. If risk is low but the consequence to the target species is high, risk management priority would be high. If, however, the risk of introducing ranavirus is high but the consequences are low, risk management priority would be low. If the IRA indicates that the consequences are high, then the recommendations to management would focus on the CCPs and how to increase the likelihood of detecting and eliminating an infected consignment in a cost effective manner.

Effective communication is required among stakeholders, both when collecting information to feed into the IRA and in terms of informing end users of the findings, management options, and their implementation. Risk communication is often centered at government level, but individual organizations such as fish farmers or herpetological societies can investigate and implement their own quarantine and surveillance guidelines with qualified diagnostic support. Cooperation and awareness at all levels will greatly reduce the risk of introducing ranavirus into an uninfected area.

Many of the facts needed to carry out a comprehensive IRA may already be available in the published scientific literature and should be used to substantiate the recommendation for a risk analysis. It is important to consider the applicability and quality of the published literature before it is used in risk analyses. Published data might be from a different species, time of year, or continent. If published data do not exist for your species or region, a pilot study can be performed to generate data. Alternatively, obtaining expert opinion following the Delphi method can be an approach to secure preliminary estimates for use in the risk analysis (Helmer 1967; Vose 2000). We recommend that all organizations that are interested in performing an IRA consult experts that study ranaviruses. The GRC is a collection of scientists, veterinarians and practitioners that can provide guidance with setting up IRAs. Each continent has a regional GRC representative that can assist or make necessary connections with experts in your region.

**Acknowledgments** We graciously thank Jason Rohr (University of South Florida) and Cheryl Briggs (University of California-Santa Barbara) for reviewing earlier drafts of our chapter and providing helpful comments. J.E. Earl's contribution was conducted while a Postdoctoral Fellow at the National Institute for Mathematical and Biological Synthesis, an Institute sponsored by the National Science Foundation, the US Department of Homeland Security, and the US Department of Agriculture through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville.

Open Access publication was made possible through grants provided by the University of Tennessee (Institute of Agriculture, Office of Research and Engagement, and Department of

Forestry, Wildlife and Fisheries), Washington State University Libraries, Gordon State College (Office of Academic Affairs), the Association of Reptilian and Amphibian Veterinarians, and the Amphibian and Reptile Conservancy.

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Appendix 1

The links below are for a guide to using the statistical program R to estimate confidence intervals, perform chi-square analyses and logistic regression, and plot the appropriate graphs.

[http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/SampleCode\\_8.5.html](http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/SampleCode_8.5.html)

[http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/SampleCode\\_8.5.R](http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/SampleCode_8.5.R)

<http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/Data.csv>

## Appendix 2

The links below are for downloading MatLab programs for doing SIR and stage-structured model simulations.

SIR model:

[http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/Example\\_SIR\\_Model.m](http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/Example_SIR_Model.m)

[http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/Example\\_SIR\\_Model.txt](http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/Example_SIR_Model.txt)

<http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/SIR.m>

<http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/SIR.txt>

Population model:

<http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/PopulationModelDetails.pdf>

<http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/PopulationModelCode.txt>

<http://fwf.ag.utk.edu/mgray/RanavirusBook/Chap8/PopulationExampleModel.m>

## References

- Allen LJS (2006) An introduction to mathematical biology. Pearson, New York
- Allender MC, Fry MM, Irizarry AR et al (2006) Intracytoplasmic inclusions in circulating leukocytes from an eastern box turtle (*Terrapene carolina carolina*) with iridoviral infection. J Wildl Dis 42:677–684
- Amos KH (1985) Procedures for the detection and identification of certain fish pathogens, 3rd edn. American Fisheries Society, Corvallis
- Amstrup SC, McDonald TL, Manly BFJ (2005) Handbook of capture-recapture analysis. Princeton University Press, Princeton

- Augusiak J, Van Den Brink PJ, Grimm V (2014) Merging validation and evaluation of ecological models to “evaluation”: a review of terminology and a practical approach. *Ecol Model* 280: 117–128
- Biek R, Funk WC, Maxwell BA et al (2002) What is missing in amphibian decline research: insights from ecological sensitivity analysis. *Conserv Biol* 16:728–734
- Brenes R (2013) Mechanisms contributing to the emergence of ranavirus in ectothermic vertebrate communities. Ph.D. Dissertation, University of Tennessee
- Brenes R, Gray MJ, Waltzek TB et al (2014a) Transmission of ranavirus between ectothermic vertebrate hosts. *PLoS One* 9:e92476
- Brenes R, Miller DL, Waltzek TB et al (2014b) Susceptibility of fish and turtles to three ranaviruses isolated from different ectothermic vertebrate classes. *J Aquat Anim Health* 26(2): 118–126
- Briggs CJ, Vredenburg VT, Knapp RA et al (2005) Investigating the population-level effects of chytridiomycosis: an emerging infectious disease of amphibians. *Ecology* 86:3149–3159
- Briggs CJ, Knapp RA, Vredenburg VT (2010) Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. *Proc Natl Acad Sci U S A* 107:9695–9700
- Brown LD, Cal TT, Dasgupta A (2001) Interval estimation for a binomial proportion. *Stat Sci* 16:101–117
- Brunner JL, Schock DM, Davidson EW et al (2004) Intraspecific reservoirs: complex life history and the persistence of a lethal ranavirus. *Ecology* 85:560–566
- Brunner JL, Schock DM, Collins JP (2007) Transmission dynamics of the amphibian ranavirus *Ambystoma tigrinum* virus. *Dis Aquat Organ* 77:87–95
- Brunner JL, Storfer A, Gray MJ, Hoverman JT (2015) Ranavirus ecology and evolution: from epidemiology to extinction. In: Gray MJ, Chinchir VG (eds) *Ranaviruses: lethal pathogens of ectothermic vertebrates*. Springer, New York
- Cariboni J, Gatelli D, Liska R et al (2007) The role of sensitivity analysis in ecological modelling. *Ecol Model* 203:167–182
- Caswell H (2000) *Matrix population models: construction, analysis, and interpretation*, 2nd edn. Sinauer, Sunderland
- Cooch EG, Conn PB, Ellner SP et al (2012) Disease dynamics in wild populations: modeling and estimation: a review. *J Ornithol* 152(2):485–509
- Dobson A (2004) Population dynamics of pathogens with multiple host species. *Am Nat* 164:S64–S78
- Dohoo I, Martin W, Stryhn H (2003) *Veterinary epidemiologic research*. AVC, Charlottetown
- Duffus ALJ (2009) Ranavirus ecology in common frogs (*Rana temporaria*) from the United Kingdom: transmission dynamics, alternate hosts, and host-strain interactions. Ph.D., University of London
- Duffus ALJ, Waltzek TB, Stöhr AC, Allender MC, Gotesman M, Whittington RJ, Hick P, Hines MK, Marschang RE (2015) Distribution and host range of ranaviruses. In: Gray MJ, Chinchir VG (eds) *Ranaviruses: lethal pathogens of ectothermic vertebrates*. Springer, New York
- Earl JE, Gray MJ (2014) Introduction of ranavirus to isolated wood frog population could cause local extinction. *EcoHealth* 11:581–592
- Evira (2013) Risk assessment of new and emerging systemic iridoviral diseases for European fish and aquatic ecosystems (RANA). <http://www.evira.fi/portal/en/about+evira/about+us/operation+areas/scientific+research/projects/previous/risk+assessment+of+new+and+emerging+systemic+iridoviral+diseases+for+european+fish+and+aquatic+ecosystems++rana+/>. Accessed 21 May 2014
- Federico P, Gross LJ, Lenhart S et al (2013) Optimal control in individual-based models: implications from aggregated methods. *Am Nat* 181:64–77
- Fenichel EP, Horan RD, Hickling GJ (2010) Management of infectious wildlife diseases: bridging conventional and bioeconomic approaches. *Ecol Appl* 20:903–914
- Forzán MJ, Wood J (2013) Low detection of ranavirus DNA in wild postmetamorphic green frogs, *Rana* (Lithobates) clamitans, despite previous or concurrent tadpole mortality. *Journal of Wildlife Diseases* 49:879–886

- Gahl MK, Calhoun AJK (2008) Landscape setting and risk of ranavirus mortality events. *Biol Conserv* 141:2679–2689
- Gotelli NJ, Ellison AM (2004) A primer of ecological statistics. Sinauer, Sunderland
- Gray MJ, Miller DL (2013) The rise of ranavirus: an emerging pathogen threatens ectothermic vertebrates. *Wildl Prof* 7:51–55
- Gray MJ, Miller DL, Hoverman JT (2009) Ecology and pathology of amphibian ranaviruses. *Dis Aquat Organ* 87:243–266
- Gray MJ, Miller DL, Hoverman JT (2012) Reliability of non-lethal surveillance methods for detecting ranavirus infection. *Dis Aquat Organ* 99:1–6
- Greer AL, Collins JP (2008) Habitat fragmentation as a result of biotic and abiotic factors controls pathogen transmission throughout a host population. *J Anim Ecol* 77:364–369
- Grimm V, Railsback SF (2005) Individual-based modeling and ecology. Princeton University Press, Princeton
- Grimm V, Railsback SF (2012) Pattern-oriented modelling: a “multiscale” for predictive systems ecology. *Philos Trans R Soc Lond B Biol Sci* 367:298–310
- Grimm V, Revilla E, Berger U et al (2005) Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310:987–991
- Grimm V, Augusiak J, Focks A et al (2014) Towards better modelling and decision support: documenting model development, testing, and analysis using TRACE. *Ecol Model* 280:129–139
- Haislip NA, Gray MJ, Hoverman JT et al (2011) Development and disease: how susceptibility to an emerging pathogen changes through anuran development. *PLoS One* 6:e22307
- Hamed MK, Gray MJ, Miller DL (2013) First report of ranavirus in plethodontid salamanders from the Mount Roger’s National Recreation Area, Virginia. *Herpetological Review* 44:455–457
- Hanski I (1999) Metapopulation ecology. Oxford University Press, New York
- Harp EM, Petranks JW (2006) Ranavirus in wood frogs (*Rana sylvatica*): potential sources of transmission within and between ponds. *J Wildl Dis* 42:307–318
- Harper EB, Rittenhouse TG, Semlitsch RD (2008) Demographic consequences of terrestrial habitat loss for pool-breeding amphibians: predicting extinction risks associated with inadequate size of buffer zones. *Conserv Biol* 22:1205–1215
- Hastings A (1997) Population biology: concepts and models. Springer, New York
- Helmer O (1967) Analysis of the future: the Delphi method. RAND, Santa Monica
- Hess G (1996) Disease in metapopulation models: implications for conservation. *Ecology* 77:1617–1632
- Hethcote HW (2000) The mathematics of infectious disease. *SIAM Rev* 42:599–653
- Hosmer DW, Lemeshow S, May S (2008) Applied survival analysis: regression modeling of time to event data. Wiley, Hoboken
- Hoverman JT, Gray MJ, Haislip NA et al (2011) Phylogeny, life history, and ecology contribute to differences in amphibian susceptibility to ranaviruses. *Ecohealth* 8:301–319
- Hoverman JT, Gray MJ, Miller DL et al (2012) Widespread occurrence of ranavirus in pond-breeding amphibian populations. *Ecohealth* 9:36–48
- Jager KJ, Van Dijk C, Zoccali C et al (2008) The analysis of survival data: the Kaplan-Meier method. *Kidney Int* 74:560–565
- Jancovich JK, Qin Q, Zhang Q-Y, Chinchir VG (2015a) Ranavirus replication: molecular, cellular, and immunological events. In: Gray MJ, Chinchir VG (eds) Ranaviruses: lethal pathogens of ectothermic vertebrates. Springer, New York
- Jancovich JK, Steckler N, Waltzek TB (2015b) Ranavirus taxonomy and phylogeny. In: Gray MJ, Chinchir VG (eds) Ranaviruses: lethal pathogens of ectothermic vertebrates. Springer, New York
- Johnson AJ, Pessier AP, Jacobson ER (2007) Experimental transmission and induction of ranaviral disease in western ornate box turtles (*Terrapene ornata ornata*) and red-eared sliders (*Trachemys scripta elegans*). *Vet Pathol* 44:285–297
- Keesing F, Holt RD, Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecol Lett* 9:485–498

- Kerby JL, Hart AJ, Storfer A (2011) Combined effects of virus, pesticide, and predator cue on the larval tiger salamander (*Ambystoma tigrinum*). *Ecohealth* 8:46–54
- Kolby JE, Smith KM, Berger L et al (2014) First evidence of amphibian chytrid fungus (*Batrachochytrium dendrobatidis*) and ranavirus in Hong Kong amphibian trade. *PLoS One* 9:e90750
- Lélu M, Langlais M, Poulle M et al (2013) When should a trophically and vertically transmitted parasite manipulate its intermediate host? The case of *Toxoplasma gondii*. *Proc Biol Sci* 280:20131143
- Lenhart S, Workman JT (2007) Optimal control applied to biological models. Chapman & Hall/CRC, Boca Raton
- Lloyd-Smith JO, Cross PC, Briggs CJ et al (2005) Should we expect population thresholds for wildlife disease? *Trends Ecol Evol* 20:511–519
- Marsh DM, Trenham PC (2001) Metapopulation dynamics and amphibian conservation. *Conserv Biol* 15:40–49
- Martín Andrés A, Silva Mato A, Tapia García JM et al (2004) Comparing the asymptotic power of exact tests in 2x2 tables. *Comput Stat Data Anal* 47:745–756
- McCallum H, Barlow N, Hone J (2001) How should pathogen transmission be modeled. *Trends Ecol Evol* 16:295–300
- Mideo N, Barclay VC, Chan BHK et al (2008) Understanding and predicting strain-specific patterns of pathogenesis in the rodent malaria, *Plasmodium chabaudi*. *Am Nat* 172:E214–E238
- Mideo N, Savill NJ, Chadwick W et al (2011) Causes of variation in malaria infection dynamics: insights from theory and data. *Am Nat* 178:174–188
- Miller DL, Gray MJ, Storfer A (2011) Ecopathology of ranaviruses infecting amphibians. *Viruses* 3:2351–2373
- Miller DL, Pessier AP, Hick P, Whittington RJ (2015) Comparative pathology of ranaviruses and diagnostic techniques. In: Gray MJ, Chinchir VG (eds) *Ranaviruses: lethal pathogens of ectothermic vertebrates*. Springer, New York
- Nazir J, Spengler M, Marschang RE (2012) Environmental persistence of amphibian and reptilian ranaviruses. *Dis Aquat Organ* 98:177–184
- OIE (2014) Aquatic Animal Health Code (online access). Office International des Epizooties, Paris. <http://www.oie.int/international-standard-setting/aquatic-code/>. Accessed 21 May 2014
- Otto SP, Day T (2007) A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press, Princeton
- Price, SJ, Garner TWJ, Nichols RA, et al. (2014) Collapse of amphibian communities due to an introduced Ranavirus. *Curr Biol* 24:2586–2591. [http://www.cell.com/current-biology/pdfExtended/S0960-9822\(14\)01149-X](http://www.cell.com/current-biology/pdfExtended/S0960-9822(14)01149-X)
- Railsback SF, Grimm V (2011) Agent-based and individual-based modeling: a practical introduction. Princeton University Press, Princeton
- Robert J, George E, De Jesús Andino F et al (2011) Waterborne infectivity of the Ranavirus frog virus 3 in *Xenopus laevis*. *Virology* 417:410–417
- Rödger D, Kielgast J, Bielby J et al (2009) Global amphibian extinction risk assessment for the panzootic chytrid fungus. *Diversity* 1:52–66
- Schloegel LM, Ferreira CM, James TY et al (2010) The North American bullfrog as a reservoir for the spread of *Batrachochytrium dendrobatidis* in Brazil. *Anim Conserv* 14:53–61
- Smith MA, Green DM (2005) Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? *Ecography* 28:110–128
- Smith KF, Behrens M, Schoegel LM et al (2009) Reducing the risk of wildlife trade. *Science* 324:594–595
- Smith SA, Seigel RA, Driscoll CP et al (2014) Detecting the extent of mortality events from ranavirus in amphibians of the Northeastern U.S. [http://rcngroups.org/sites/default/files/original\\_proposals/RCN%202012\(1\)%20Ranavirus%20in%20amphibians.pdf](http://rcngroups.org/sites/default/files/original_proposals/RCN%202012(1)%20Ranavirus%20in%20amphibians.pdf). Accessed 21 May 2014
- St-Amour V, Lesbarrères D (2007) Detecting Ranavirus in toe clips: an alternative to lethal sampling methods. *Conserv Genet* 8:1247–1250
- Storfer A, Alfaro ME, Ridenhour BJ et al (2007) Phylogenetic concordance analysis shows an emerging pathogen is novel and endemic. *Ecol Lett* 10:1075–1083

- Sutton WB, Gray MJ, Hoverman JT et al (2014) Trends in ranavirus prevalence among plethodontid salamanders in the Great Smoky Mountains National Park. *EcoHealth* DOI:[10.1007/s10393-014-0994-z](https://doi.org/10.1007/s10393-014-0994-z)
- Teacher AGF, Cunningham AA, Garner TWJ (2010) Assessing the long-term impact of *Ranavirus* infection in wild common frog populations. *Anim Conserv* 13:514–522
- Thoesen JC (1994) Blue book: suggested procedures for the detection and identification of certain finfish and shellfish pathogens. Version 1. American Fisheries Society, Bethesda
- Thomson DL, Cooch EG, Conroy MJ (2009) Modeling demographic processes in marked populations. Environmental and ecological statistics. Springer, New York
- Todd-Thompson M (2010) Seasonality, variation in species prevalence, and localized disease for ranavirus in cades cove (Great Smoky Mountains National Park) amphibians. M.S., University of Tennessee. [http://trace.tennessee.edu/utk\\_gradthes/665/](http://trace.tennessee.edu/utk_gradthes/665/). Accessed 21 May 2014
- Vose D (2000) Risk analysis—a quantitative guide, 2nd edn. Wiley, Chichester
- Waltzek TB, Miller DL, Gray MJ et al (2014) New disease records for hatchery-reared sturgeon. I. Expansion of Frog Virus 3 into *Scaphirhynchus albus*. *Diseases of Aquatic Organisms* 111:219–227
- Wells KD (2007) The ecology and behavior of amphibians. University of Chicago Press, Chicago
- Wiegand T, Jeltsch F, Hanski I et al (2003) Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. *Oikos* 100:209–222
- Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic, San Diego
- Wilson EB (1927) Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc* 22:209–212
- Wobeser GA (2006) Essentials of disease in wild animals. Blackwell, Ames
- Woodhams DC, Alford RA, Briggs CJ et al (2008) Life-history trade-offs influence disease in changing climates: strategies of an amphibian pathogen. *Ecology* 89:1627–1639
- Woodhams DC, Bosch J, Briggs CJ et al (2011) Mitigating amphibian disease: strategies to maintain wild populations and control chytridiomycosis. *Front Zool* 8:8
- Zuur AF, Savaliev AA, Ieno EN (2012) Zero inflated models and generalized linear mixed models with R. Highland Statistics, Newburgh